Evaluating the Historical/Cohort approach for building the consideration set in route choice modeling using smart card data from a large-scale public transport network

Jacqueline Arriagada[1a], C. Angelo Guevara[bc], Marcela A. Munizaga[bc]

a. Departamento de Ingeniería Civil Industrial, Universidad de Chile, Santiago, Chile.
b. Departamento de Ingeniería Civil, Universidad de Chile, Santiago, Chile
c. Instituto Sistemas Complejos de Ingeniería (ISCI)

Abstract

This study contributes to the understanding of public transport passenger route choice modeling by evaluating the consideration set built with historical route choices of passengers in comparison with other consideration set generation methods widely used in public transport route choice literature. We define the Historical/Cohort approach as building a practical consideration set from observed choices that occurred in some past instances of the traveler or observed choices of other users in the same cohort in cross-section data. Villalobos and Guevara (2021) evidenced that the Historical/Cohort approach was the only one, among others analyzed approaches, that was able to recover the population parameters in the experiments. We extend this result by providing a formal theoretical demonstration that, under appropriate assumptions, shows that the Historical/Cohort approach to building the consideration set can obtain consistent estimators of the population parameters. Using smart card data of the Santiago public transport system, we assess the relative performance of the Historical/Cohort approach, with respect to a set of other practical methods to address the consideration set problem. To evaluate the prediction and estimation performance, we estimated multinomial and path size logit models. In line with the theoretical demonstration, the results show that the Cohort/Historical approach allows estimating models with better prediction abilities for the choices of passengers in the prediction sample.

Keywords: public transport, route choice, smart card data, passenger behavior

1. INTRODUCTION

One of the biggest challenges when estimating route choice models is to identify the complete set of alternative routes considered by the passengers when they want to travel between an origin and a destination. This set is called the consideration set and it plays a fundamental role in route choice modeling, since some studies have shown that the composition of the consideration set influences the route choice model estimates and the choice probabilities (Bliemer & Bovy, 2008; Prato & Bekhor, 2007). However, little is known about the impacts of the composition of the consideration set in public transport modeling. This shortfall applies both to the estimation and to the predictive performance of route choice models estimated with different compositions of the consideration set. Our study contributes to shed some light into this subject.

The literature identifies three types of approaches to handle with the unobserved choice set problem. First, there are theoretical contributions related to discrete choice models that explicitly include the construction of the consideration-set in their analysis. The seminal work to this regard

[1] jarriagadafe@ing.uchile.cl

corresponds to Manski (1977) who proposes an approach where the consideration-set is treated as a latent variable. Even though Manski (1977)'s approach theoretically solves the problem, in practical terms it is extremely difficult to appropriately define a practical function for the probability of considering a given set and it involves great computing costs in enumerating large number of choice sets, especially in the case of dense transit networks. To solve the first practical problem, Swait & Ben-Akiva (1987) and Ben-Akiva & Boccara (1995) propose using individual characteristics (e.g., income or the driver's license ownership) or restrictions (e.g., distance) to develop expressions for the consideration-set probability. Although this approach is intuitively appealing, it should be recognized that these definitions are still ad-hoc solutions to the problem.

Second, there are empirical contributions related to the consideration-set, mainly in marketing, in which several factors of this problem are studied, such as size and potential determination attributes. In the marketing research field, the consideration-set has been profusely studied in the context of the commercialization of different types of products, since it is key for a brand to be included in the consumer's consideration-set to be chosen. For example, Brown & Wildt (1992) investigated the size of the consideration-set concluding that a consumer's consideration-set includes, on average, 4.54 alternatives, in the context of fast food, soft drinks, and gasoline stations. Hauser (2014) used similar methods to study different types of heuristics that consumers could use to generate their consideration-set and suggested ways to identify them. Crawford et al. (2021) introduced the sufficient set as a method to identify subsets of true choice sets through the observed choices and they showed that using the sufficient set as consideration set, the preferences' parameters can be consistently estimated. To this regard, Villalobos & Guevara (2021) show results that suggest that using stated preferences to collect information on the consideration-set may be prone to severe hypothetical bias, reflected in the fact that the size of the consideration-set gathered from such tools depends systematically on the experimental setting.

A third approach to the consideration-set problem, mainly on route choice modeling, is aimed to propose practical methods for generating the consideration-set. A practical method is usually an algorithm or heuristic that try to reproduce the behavioral passengers to choose a route. Our research focuses on this viewpoint, with an emphasis on the case of public transportation.

The existing literature on route preferences of public transport users have used both stated preference (SP) data and revealed preference (RP) data. The studies that have used SP data (Eluru, Chakour, & El-Geneidy, 2012; Grison, Burkhardt, & Gyselinck, 2017; Vrtic & Axhausen, 2002) obtain the information from a survey that asks people to choose a route alternative from a set of alternatives in a hypothetical route choice situation, that may or may not be pivoted in a real trip situation. This methodology to obtain the data has the advantage that it is relatively inexpensive, and that the researcher knows the consideration set faced by the respondent, but it introduces hypothetical bias, since the user does not experience the actual trip.

On the other hand, the use of RP data has the advantage that it reflects information about the choices of users in real situations, but in this case the researcher does not know the true consideration set and collection costs are more expensive. Usually, public transport route choice studies that work with RP data use traditional travel survey methods to capture the consideration set which are not only expensive, but also impractical for large samples (Z. Guo, 2011; Z. Guo & Wilson, 2011; S Raveau & Muñoz, 2014; Sebastián Raveau, Guo, Muñoz, & Wilson, 2014; Sebastián Raveau, Muñoz, & de Grange, 2011; Ton et al., 2020; Vrtic & Axhausen, 2002). Recently, some authors have dealt with this problem using smart card data, which has as main purpose to collect public transport revenue and, as side benefit, allows to collect a large quantity of very detailed data about choices of public transport users at relatively lower costs and practical limitations, as well as unprecedent granularity (Pelletier et al., 2011).

Another big challenge in the public transport route choice modelling is the high level of correlation between routes alternatives that share a large number of links (Prato, 2009), which alters the choice probabilities of overlapping routes. In this line, it is necessary that route choice models represent properly the correlation structure among alternative routes. Note that most studies on public transport passengers' route choice behavior have used the Multinomial Logit (MNL) discrete choice model (Grison et al., 2017; Guo, 2011; Jánošíková et al., 2014; Nassir et al., 2018; Raveau & Muñoz, 2014; Raveau et al., 2014; Raveau et al., 2011; Vrtic & Axhausen, 2002; Kim et al., 2020), which assumes independence of alternatives, and therefore the correlation problem due to overlapping route segments is ignored in this case. To address this limitation, the analytical approach of Path Size (PS) Logit models have often been adopted, which accounts for the correlation by adding a deterministic term that reduces the utility of overlapped alternatives (Anderson et al., 2017; Bovy & Hoogendoorn-Lanser, 2005; De Grange et al., 2012; Hoogendoorn-Lanser et al., 2005; Nielsen et al., 2021; Tan et al., 2015; Yap et al., 2020). In this work we implement both type of models, the basic one (MNL model), and the PS Logit model to capture the correlation among alternative routes.

In the context of public transport route choice studies that use RP data, the consideration set has been built in practice by identifying route alternatives that appear to be logical, feasible, or attractive to the passengers, under the eyes of the modeler. This consideration set generation process is a complex task, since, usually, there are countless feasible alternative routes in a transport network, especially in a dense and multimodal one. Additionally, there is evidence that suggests that the number of route alternatives known and considered by the passenger is substantially smaller than the available ones (Hoogendoorn-Lanser & Van Nes, 2004). Two alternatives for identifying the consideration set in practice can be distinguished in the applied literature: built it using some algorithm or heuristic that emulates how individuals may build the consideration set or to impute it form historical data.

Most of the heuristics used to build feasible consideration sets in practice are based on repeatedly varying some type of deterministic shortest path. The **K-shortest path approach** is based in the behavioral assumption that users consider a limited number of alternative routes, and consequently, it finds the best k paths using some link additive generalized cost (Prato, 2009). A well-known K-shortest path algorithm was proposed by Yen (1971). It consists in two main parts: first the shortest path is found using a traditional shortest path algorithm, then the subsequent shortest paths are found blocking an intermediate link of a previously found path to force the algorithm to find another path. This method has been mostly used for private transportation modeling. To the best of our knowledge, there is only one study on public transport route choice that have used the k-shortest path approach (Tan et al., 2015). The **Labeling approach** is based in the behavioral assumption that users consider different objectives to select the considered route alternatives, then each label correspond to a different objective function for which the path is optimum (Ben-Akiva, M., Bergman, M. J., Daly, A. J., & Ramaswamy, 1984). A number of studies on public transport route choice have used the Labeling approach (Z. Guo, 2011; Z. Guo & Wilson, 2011; Tan et al., 2015). The **Link elimination approach** searches repetitively the shortest path after removing a part or all shortest path links from the previous optimum path. We found only one study on public transport route choice that have used the Link elimination approach (Tan et al., 2015). The **Link penalty approach** also searches repetitively the shortest path, but instead of removing links, it adds a penalty on the impedance of all links in the optimal path. We again found only one study on public transport route choice that have used the Link penalty approach (Tan et al., 2015). Other group of path generation methods are the **stochastic path methods**, which can be classified into simulation and double stochastic approaches. They are also based on repeated shortest path searches where the link impedances are drawn from a probability distribution (Prato, 2009). We found two studies on public transport route choice that have used stochastic path generation algorithms to identify the consideration set (Anderson, Nielsen, & Prato, 2017; Tan et al., 2015). Previous public transport route choice studies have explored the **branch and bound algorithm**, which is a constrained enumeration method, and it is based in the assumption that passengers choose routes using behavioral rules different to the minimum cost path. Finally, some studies have used a **combination of previous methods** for modeling public transport. For example, Fiorenzo-Catalano, Van Nes, & Bovy (2004) applied a combination of the labeling approach and the simulation approach in a multimodal network in Netherlands and Tan et al., (2015) applied a combination of the link elimination, labeling, k shortest paths, and simulation approaches in a multimodal network in Singapore.

Another practical approach used in the last decade to tackle the consideration set problem has been to impute it from historical data, i.e. previous choices made in similar situation (Jánošíkova et al., 2014; Kim et al., 2020). We denominate this method as the **Historical/Cohort approach**, which can be formally defined as building a practical consideration set from observed choices that occurred in some past instances of the traveler or observed choices of other users in the same cohort in cross section data. The cohort in cross sectional data may correspond to trips that were performed on the same OD pair, period, trip purpose, by individuals of the income group and household composition. Identifying the Historical/Cohort consideration set is difficult with traditional transport survey techniques, since the user may forget to mention an alternative, may not share some information, or may make errors in describing the alternatives (Hoogendoorn-Lanser & Van Nes, 2004). However, the introduction of smart card automated fare collection systems in public transport allows to researcher to identify the Historical/Cohort consideration set without much effort. Sebastián Raveau et al. (2014) states that "the consideration set build with the observed choices can be considered as an ideal scenario", because in this case there is no need to use behavioral assumptions as the methods presented previously. In this line, Villalobos and Guevara (2021) investigated the impact of different heuristics to construct the consideration set in the estimation of route choice models, using Monte Carlo experiments, and they evidenced that the Historical/Cohort approach was the only one of the methods analyzed that was able to recover the population parameters in the experiments.

In this paper we extend the empirical results of Villalobos and Guevara (2021) in two ways. We first provide a formal theoretical demonstration that, under appropriate assumptions, the Historical/Cohort approach to build the consideration set is able to provide consistent estimators of the population parameters. This differentiates the Historical/Cohort from other practical approaches to the consideration set problem that are inherently heuristic in nature.

Besides, we assess the relative performance of the Historical/Cohort approach, with respect to a set of other practical methods to address the consideration set problem, using revealed preference data of public transportation route choices. In this line, we evaluate the effect of six consideration set generation approaches: k-shortest paths, link elimination, link penalty, labeling, simulation, and Historical/Cohort consideration set. The revealed preference data is built using smart card's transaction data of three weeks representing the actual route choice behavior of passengers that travel in the public transport system of Santiago, Chile. For the analysis, we use two specifications: multinomial logit and the path size logit models, which accounts for the correlation of routes. With this, we evaluate first the impact of different consideration set generation approaches by assessing the plausibility of the model estimates and the in-sample fit attained by each approach using different statistics and criteria. Then, we conclude studying the out-of-sample prediction performance attained by each method on a fourth week of data.

The remainder of this paper is organized as follows. The next section presents a formal demonstration of why the Historical/Cohort approach allows obtaining consistent estimators of the model parameters. Then we describe the case study, the transit network. The paper discusses the proposed methods used for the consideration set and route choice models. The model estimation and prediction results are then introduced. The last section draws conclusions and discusses policy and research implications.

2. CONDITIONS UNDER WHICH HISTORICAL/COHORT APPROACH DELIVERS CONSISTENT ESTIMATORS

The ability of the Historical/Cohort approach to recover the population parameters is theoretically proved in this section. The demonstration is based on an adaptation of the theorem of sampling of alternatives (McFadden, 1978), in which each historical/cohort choice is understood as a draw from the true consideration set, which is latent to the researcher. Under this setting we show that the sampling correction needed for estimation cancels out as the number of observations grows what explains why the method works even when the sampling correction is ignored.

Consider a Random Utility Model (RUM) setting in which the utility U_{in} that an individual n retrieves from alternative i can be written as the sum of a systematic part $V_{in}$and a random error term ε_{in} as shown in Equation (1), where V_{in} depends on attributes x_{in} and population parameters β^*.

$$U_{in} = V_{in} + \varepsilon_{in} = V(x_{in}, \beta^*) + \varepsilon_{in} \tag{1}$$

Then, if ε_{in} is distributed *iid* Extreme Value $(0, \mu)$, the probability that n will choose alternative i will correspond to the Logit model shown in Equation (2), where C_n is the true consideration-set of J_n elements from which individual n selects one alternative. The scale μ in Equation (2) is not identifiable and is thus usually normalized to equal 1 to grant identification.

$$P_n(i) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}} \tag{2}$$

Consider that the true consideration-set C_n is latent to the researcher, but that she can observe the R choices that occurred in past instances. These observations could correspond to choices of the same individual n or, assuming group homogeneity, choices made by diverse individuals that faced the same choice situation. In practice, data of this type can be gathered, for example, from a series of supermarket purchases or, in the case of a commuting mode or route, from a series of passive data records across weekdays. In cross sectional data, this information may be obtained by observing trips on the same OD pair, period, trip purpose, income group, household composition, etc. As it is described in the next section, in the case study considered in this article, based on passive smart card data, we consider that a cohort corresponds to trips made in the same OD pair, defined at the bus-stop level, and period of the day, within the previous three weeks of available data. We denominate this practical choice-set as the Historical/Cohort consideration set. Since the number of previous choices within the previous three weeks may differ by OD pair, formally R shall depend on the OD pair but, for notational simplicity, we will avoid adding an OD subscript.

The researcher is interested in modeling the choice occurring at the instance $R + 1$. For this she builds a practical consideration set that includes all the alternatives that were observed in the previous R instances, plus the alternative chosen on instance $R + 1$, if it was not already included. We assume attributes and choice-set's invariability across the $R + 1$ instances. This implies that neither the x_{in} nor the consideration-set change across time. Regarding the internal validity of the attribute's invariability assumption, it could easily hold for a supermarket purchasing example but may be more questionable for the commuter public transportation example, in which one may only expect some level of stability. The assumption of the invariability of the consideration set may most likely holds in practice for habitual choices, like commuting in the public transport example or when buying customary groceries in the supermarket example. The degree of external validity of these assumptions may only be tested in the field, with real data. This is the main purpose of our case study presented latter.

Under these assumptions, the key step toward the demonstration of the consistency of the Historical/Cohort approach resides in noting that the R previous choices could be understood as draws with replacement from the true consideration set.

Formally, the problem originally tackled by McFadden (1978) was that the true consideration set C_n was too large to be processed by the researcher in practice, what was solved by building a reduced practical set $D_n \subseteq C_n$ for estimation, using some known sampling protocol conditional on the chosen alternative. Formally, $\pi(D_n|j)$ corresponds to the conditional probability that the researcher would sample the set D_n, given that alternative j was chosen by individual n. Under this setting McFadden (1978) showed that maximizing a pseudo-loglikelihood using the choice probabilities shown in Equation (3), one can obtain consistent estimators of the model parameters.

$$\pi(i|D_n) = \frac{e^{V_{in} + ln\pi(D_n|i)}}{\sum_{j \in D_n} e^{V_{jn} + ln\pi(D_n|j)}} \tag{3}$$

The merit of the model depicted in Equation (3) is that it only depends on the alternatives of the reduced set D_n, reducing a problem of possibly millions of alternatives to just a few dozens. Besides, the resulting model has a closed form that corresponds to a simple Logit model with a correction term by alternative $ln\pi(D_n|j)$ that only depends on the sampling protocol. In some cases, this correction even cancels out across alternative and can be therefore ignored. This occurs, for example, when the protocol used to build D_n corresponds to draw the chosen alternative and then to add a given number of nonchosen alternatives randomly drawn from C_n. This result holds thanks to the IIA property of the Logit model, but was extended to more flexible models like GEV (MEV), Logit Mixture and RRM, by Guevara & Ben-Akiva (2013a, b) and Guevara et al. (2016), respectively.

As stated before, the Historical/Chohort approach to the consideration set problem can be seen as a problem of estimation with importance sampling of alternatives with replacement, feature that was studied by Ben-Akiva & Lerman (1985) and revisited by Ben-Akiva (1989). Using Ben-Akiva (1989) result, it can be shown that, for the Historical/Cohort approach, the sampling correction $\pi(D_n|i)$ corresponds to the expression shown in Equation (4), where n_j is the number of times alternative j was chosen in the $R+1$ instances.

$$\pi(D_n|i) = \frac{n_i}{P_n(i)} \frac{R!}{\prod_{j\epsilon D} n_j!} \prod_{j\epsilon D} P_n(j)^{n_j} = \frac{n_i}{P_n(i)} K_D \approx \frac{n_i}{n_i/_R} K_D = RK_D \tag{4}$$

The demonstration continues by noting first that the term K_D that does not change across alternatives and that, as the number of instances R grow, the choice probability $P_n(i)$ will become closer to $n_i/_R$, resulting in that the sampling correction approximately becomes RK_D, a term that does not depend on the alternative. Since the sampling correction in Equation (4) enters the choice probability shown in Equation (3), this means that the correction cancels out and can then be simply ignored in the likelihood. This provides a theoretical support for the empirical results reported by Villalobos and Guevara (2021). Based on this demonstration, it can be affirmed that, under the invariability assumption, the Historical/Cohort approach allows obtaining the same or better performance estimation and prediction results than other heuristic methods commonly used to identify the consideration set. The case study revised in the next section is aimed at studying this hypothesis.

3. CASE STUDY: DATA SOURCES AND RESEARCH METHODOLOGY

In this section we describe the implementation of the case study, which consists of i) description of the available data set, ii) definition of the urban modeling network, iii) application of six consideration set generation techniques to the urban network of Santiago, Chile; iv) estimation of the route choice models using the consideration sets generated in (ii), and (v) evaluation of the model performance for the different consideration set.

3.1 Smart Card Data on Route Choice from Santiago, Chile

The analysis for this study was carried out using the Santiago, Chile multimodal public transport network, known as Transantiago. Santiago is the capital of Chile and has a population of over 7 million inhabitants. The public transport system serves roughly 50% of motorized trips and it is operated by headway scheduling; therefore, lines do not have fixed-time schedules. The fare system is fully integrated, with an almost flat fare between urban buses, Metro, and some rail services, allowing up to three trip legs within a two-hour time window. In a typical week, 3 million cards (passengers) use the system to make 25.5 million trips. The network includes 7 metro lines, more than 300 bi-directional transit lines, and one rail service.

Very detailed demand information was obtained from the automatic fare collection (AFC) system in Transantiago (Gschwender et al., 2016). The smart card *bip!* is the only accepted payment method. Passengers must validate when boarding a bus or entering a Metro station. No alighting validation is required for bus or Metro trips. Around 27% of passengers evade fare payment on buses. Bus stops with particularly high demand have an off-vehicle payment system called *zona paga* (payment zone), where passengers validate when they enter the bus stop area and then board any bus without further validation. The data is already processed to estimate the boarding and alighting position for all validations and the trips (stages) associated with an origin-destination journey using the Munizaga & Palma (2012) methodology.

We selected trips made in the public transport system during May 2018. Specifically, we use 15 weekdays (three weeks) to estimate models and 5 weekdays (last week) to evaluate the prediction accuracy of the models.

3.2 Network representation

For representing the urban transit network, we use the network representation proposed by Spiess and Florian (1989) and Cepeda et al. (2006), which is a frequency-based network formulation, through a direct graph $G = (N, A)$, where N represents the nodes of the network and A represents the arcs of the network. This representation contains two subsets of nodes, the stop nodes ($N_s \subseteq N$) are used to represent the bus stops, metro station, or train station line nodes ($N_l \subseteq N$) are used to represent the transit lines (bus, metro, or train). All nodes that represent the same transit line are connected by an on-board arc. When a transit line, represented by a line-node A, serves a bus stop, represented by a stop-node B, both nodes, A and B, are connected by a boarding arc and an alighting arc. When two stop nodes are separated by a walkable distance, they are connected by a walking arc. Additionally, each arc $a \in A$ is characterized by t_a, which is a nonnegative travel time and f_a, which is a nonnegative frequency. This frequency is used to estimate the waiting time, and it is assumed that the arrivals of the different transit lines are independent and exponentially distributed. This assumption has been adopted by several authors (Spiess and Florian, 1989; Cepeda et al., 2006), since the exponential distribution (or the gamma, which is the sum of exponentials) has been found to fit real data well (Guo et al., 2011). Alighting, walking, and on-board arcs do not have waiting time, then they are assigned infinite frequencies.

Therefore, waiting time in the boarding links is obtained assuming a poisson process, yielding an average value of one over the observed frequency of the transit line. The observed frequency is obtained from the Automatic Vehicle Location (AVL) data. Travel time in-vehicle in the on-board links is obtained from a combination of the AVL data and General Transit Feed Specification (GTFS) data. For a link which involves a transit line with several expeditions, the process considers all expeditions of the transit lines and take an average of the travel time in the link. The travel time of the link per expedition is obtained from the AVL data, and if it is not registered it is obtained from the GTFS data. Since the AVL data it is not available for the metro or the train, the travel time in a link of these transport modes is obtained from the operational information from the operator company. The walking time in walking links is obtained assuming that passengers walk with a speed of 4 km/hour, which is a standard value within transportation studies, and considering the Manhattan distance between nodes, since some studies have found that it approximates better the network distance than the Euclidean distance (Mora-Garcia, Marti-Ciriquian, Perez-Sanchez, & Cespedes-Lopez, 2018; Tien, MacDonald, & Xu, 2011). The total network used in this study contains 46,583 nodes and 117,816 links.

3.3 Six approaches for generating the consideration set

Six techniques for representing the consideration set are applied to the urban transit network of Santiago, Chile. Prato (2009) provide an explanation about the Labeling, the Link elimination, the Link penalty, the the K-shortest paths, and simulation approaches, which are used in this study. The Historical/Cohort consideration set approach is explained in section 2 of this paper.

The labeling approach using different criterion or cost functions. In this study this is applied by minimizing the cost path using 6 cost functions, called labels. Label 1 to 3 use only one attribute path in the cost function and label 4 to 6 weights multiple paths attributes. As an example, label 1, which is the minimal travel time in-vehicle, generates the alternative route with the minimal travel time in-vehicle between the origin and destination zones. Label 4 assumes that passengers perceive travel time in-vehicle, waiting time and walking transfer time equally. Label 6 penalizes waiting time and walking transfer time, since several authors have shown that passengers have different perception for different route attributes, specifically they have found that waiting time and walking transfer time are more burdensome than in-vehicle travel time (Nassir, Hickman, & Ma, 2018; S Raveau & Muñoz, 2014; Tan et al., 2015). In this study we use the penalization 1.6 for waiting time and 3 for walking transfer time, since Arriagada et al. (2021) found these mean factors with Path size logit models using smart card data from Santiago, Chile. Label 7 use the penalization of the label 6 and penalizes transfers bus-to-bus, metro-to-bus, and bus-to-metro. We penalize each transfer with 13 minutes of travel time, which was found by Arriagada et al. (2021).

The link elimination approach searches for the shortest path repetitively after removing a link from an optimal path. This approach follows the stages: (a) identification of the shortest travel time path, (b) elimination from the shortest path the closer link to the origin that has not been removed previously, and (c) identification of the shortest travel time path. We followed the procedure used by Rui (2016), however there are other ways to implement this approach (see e.g. Prato & Bekhor, 2007). It must be noted that when all links along the first shortest path have been eliminated, the iteration will move to the next generated path. The procedure ends when a maximum number of alternative paths (N) is reached. In this work we use N=20 since the maximum size of observed paths from smart card data is 18. The cost consists of in-vehicle travel time, waiting time, walking time, and transfer penalty.

The link penalty approach also searches for the shortest path repetitively, but unlike to the previous method imposes a penalty on the cost of all links the optimal path instead of removing a link. This approach follows the stages: (a) identification of the shortest travel time path, (b) penalization the shortest path links with a factor of 1.2, and (c) identification of the shortest travel time path. We followed the procedure used by Prato & Bekhor (2007). The procedure ends when a maximum number of alternative paths (N) is reached. In this work we use N=20 since the maximum size of observed paths from smart card data is 18. The cost consists of in-vehicle travel time, waiting time, walking time, and transfer penalty.

The K shortest path approach consists in the identification of the best K paths according to the link cost function. The behavioral assumption is that the passenger makes choices from a limit size consideration set, avoiding costly alternative. In this study we use the algorithm proposed by Yen (1971) and K is set to 20 since the maximum size of observed paths from smart card data is 18. The cost consists of in-vehicle travel time, waiting time, walking time, and transfer penalty.

The simulation approach searches for the shortest path for each random draw of link cost function from a truncated normal distribution with the mean equal to the original cost of the link and the standard deviation equal to the 20% of the original value. 50 draws of randomized cost function were performed for each OD pair. The cost function includes in-vehicle travel time, waiting time, walking time, and transfer penalty.

The Historical/Cohort approach consists of identifying all alternatives recorded and observed in the smart card data for each OD pair in the past by any traveler. Each alternative is characterized by the stops and transit lines used by the passengers that have travelled in the OD pair. The intuition of this approach is that all travelers in the same OD pair might share the same consideration set and thus, the historical choices made by the individuals from the same cohort (OD pair) will necessarily belong to the true consideration set. This approach has been used by (Jánošíkova et al., 2014; Kim et al., 2020).

Please note that all approaches, except the Historical/Cohort, require the specification of a transit network and they can generate alternative paths that contain walking stages at the beginning and/or at the end. Alternatives of this style cannot be observed in the smart card data, therefore we imposed on all heuristic approaches that the first link should not be a walking link.

3.4 Specification of route choice models

In this study we use two types of RUM models, the Multinomial Logit (MNL) discrete choice model and the Path Size Logit (PSL) model. The MNL model is the basic one, which has been used for most studies on public transport passengers' route choice behavior (Grison et al., 2017; Guo, 2011; Jánošíková et al., 2014; Nassir et al., 2018; Raveau & Muñoz, 2014; Raveau et al., 2014; Raveau et al., 2011; Vrtic & Axhausen, 2002; Ton et al., 2020; Kim et al., 2019). Since, route choice models presents a correlation between alternatives due to overlapping route segments, it is necessary to correct the MNL model, which assumes independences of alternatives. To address this problem, the analytical approach of PSL models have often been adopted, which accounts for the correlation by adding a deterministic term that reduces the utility of overlapped alternatives (Anderson et al., 2017; Bovy & Hoogendoorn-Lanser, 2005; De Grange et al., 2012; Hoogendoorn-Lanser et al., 2005; Nielsen et al., 2021; Tan et al., 2015).

The deterministic component of the MNL model is specified in Equation (5), where i represents the alternative route, TT_i is on-board travel time, IWT_i is the initial waiting time, TWT_i is the transfer waiting time, $TWalT_i$ is the transfer walking time, TM_i is the number of transfer between different lines of metro, and OT_i is the number of transfers of type: metro-to-bus, bus-to-metro or bus-to-bus. On the other hand, the PSL model is presented in the Equation (6), which contains all attributes introduced in the MNL model and adds the path size correction (PSC_i) term to capture the correlation due to overlapping between alternative routes. Path size correction introduces a negative factor that decreases the deterministic utility of alternative routes that have correlation with other routes. We have used the expression according to Bovy et al. (2008) in Equation (7), where L_r is the length of the route section r, L_i is the length of route i, ζ_i is the set of route sections belonging to route i, and δ_{rk} is the route section-route incidence number, which takes a value of 1 if route k uses route section r and a value of 0 otherwise.

$$V_i = \beta_{TT}TT_i + \beta_{IWT}IWT_i + \beta_{TWT}TWT_i + \beta_{TWalT}TWalT_i + \beta_{TM}TM_i + \beta_{OT}OT_i \quad (5)$$

$$V_i = \beta_{TT}TT_i + \beta_{IWT}IWT_i + \beta_{TWT}TWT_i + \beta_{TWalT}TWalT_i + \beta_{TM}TM_i + \beta_{OT}OT_i + \beta_{PSC}PSC_i \quad (6)$$

$$PSC_i = \sum_{r \in \zeta_i} \left(\frac{L_r}{L_i} \ln \frac{1}{\sum_{k \in C_p} \delta_{rk}} \right) \quad (7)$$

Because of the Gumbel distribution of the error term, the probability of passenger n choosing alternative i given consideration set C is expressed as in the Equation (2). The closed-form logit formula of the Logit model allows for a simple estimation of the fixed coefficients by maximizing the likelihood function.

3.5 Evaluation of methods

The purpose of the evaluation process is to analyze the performance of the different consideration set generation approaches in a qualitatively and a quantitatively perspective. Since the route choice models constructed with different consideration set generation approaches are not comparable between them using statistics with likelihood, we use the Cross-validation method, which consists of evaluating the prediction performance of the models for a part of the data set that was not used for the model estimation (Shao, 1993). In our case the cross-section units are period of times, since the set of observations are split in two subsamples, the first subsample and period to estimate models and the second subsample and period to evaluate the prediction of models. In validation sample we use the First Preference Recovery (FPR) index, which is the proportion of observations that use the route alternative with the highest chosen probability and the Average Likelihood (AL). The FPR is shown in the Equation (8) and the AL is shown in the Equation (9). In those equation $C_n(i) = 1$ if path i is chosen by observation n, 0 otherwise, N is the number of observations, and $P_n(i)$ is the calculated probability of observation n choosing path i.

$$FPR = \frac{\sum_n \sum_i 1 * C_n(i)}{N} \quad (8)$$

$$AL = \frac{\sum_n \sum_i P_n(i) * C_n(i)}{N} \quad (9)$$

4. RESULTS

4.1 Consideration set generation approaches

The objective of a consideration set generation technique is to reproduce the actual behavior of passengers and therefore to get the maximum percentage of observations for which a path generation approach reproduces the actual behavior (Prato & Bekhor, 2003). In this work we use three coverage indicators: the Trip Coverage (TP), which is the percentage of trips which the chosen alternative is available in the generated consideration set. It represents the percentage of trips that can be modeled with each approach and indicates the effectiveness of the consideration set approach; the Efficient Coverage (EC), which is the percentage of the generated alternatives being an observed alternative. This indicates the efficiency level of the consideration set approach to produce the paths used by passengers and not to produce paths do not use by the passengers; the Passenger Path Coverage (PPC), which is the percentage of observed paths being available in the generated consideration set. This indicates the comprehensiveness level of the generated consideration set with the observed alternatives. The EC and the PPC were proposed by Tan (2016). It is important to note that the trip coverage is the most important indicator since it affects the second stage of a route choice model. ***Table 1*** shows the three coverage indicators for each consideration set generation approach in the estimation sample and in the prediction sample. Please note that the Historical/Cohort approach is constructed with the observed paths in the estimation sample, then in this period all its coverage indicators are equal to one, however in the prediction sample the observed paths can be different to the estimation paths, consequently not necessarily all coverage indicators are equal to one.

As it can be seen in ***Table 1***, the Historical/Cohort approach in the prediction sample reaches the highest coverage indicators among all heuristics. On the other hand, out of the five heuristic approaches, the Labeling approach obtained the highest efficient coverage, which means that it is the best in generating the paths that are mostly chosen by the passengers. This finding was also reported by Rui (2016). The investigation about the uncovered alternatives in the observed trips shows that, in general, these paths have similar costs to the path generated by the labels that minimize the total travel time; however, they are not captured by any label. While the generated alternatives used by passengers are captured mainly by the label that minimizes the total travel time, the label that minimizes the in-vehicle travel time generates paths with many transfers, and label that minimizes the waiting time, and the label that minimizes the number of transfers generate alternatives with long in-vehicle travel time.

If we now turn to the simulation approach, it can be observed that it obtains the poorest performance in the TC and the PPC, even though the number of draws (50) is higher than the K value for the k-shortest path approach and the N value for the link elimination, and the link penalty approaches. The low level of the passenger path coverage and trip coverage means that the simulation approach generated paths with the lowest coverage of observed paths. These problems occur because many non-attractive paths are generated, specifically with more transfers than the observed paths. Additionally, the simulation, the link penalty, and the link elimination approaches obtained the lowest level of EC, which means that they generate more paths not used by the passengers than the other approaches. In contrast to the results of simulation approach performance, the link penalty and link elimination approaches are the most effective to produce the observed alternatives since they obtained the highest TC and PPC. Finally, the k-shortest path approach obtains poorer performance than the link penalty and link elimination approaches since it obtains a lower value for the TC and the PPC.

Table 1: Trip Coverage (TC), Efficient Coverage (EC), and Passenger Path Coverage (PPC) of each consideration set generation approach

Consideration set generation approach	Estimation sample			Prediction sample		
	TP	EC	PPC	TP	EC	PPC
Historical/Cohort	1	1	1	0.98	0.80	0.94

K-shortest paths	0.76	0.53	0.64	0.77	0.52	0.71
Labeling	0.73	0.56	0.62	0.74	0.55	0.67
Link elimination	0.83	0.35	0.72	0.83	0.33	0.77
Link penalty	0.90	0.14	0.80	0.89	0.13	0.84
Simulation	0.65	0.29	0.54	0.67	0.30	0.60

In order to evaluate the composition of the choice set, we present, in
Table *2*, the average number of paths per consideration set and the average of the average of path attributes for the alternatives in each OD pair for each generated consideration set approach. As shown in
Table *2*, the labeling approach is the only one that generates in average a small size of path sets than the historical/cohort approach, suggesting that use this method alone might results in inaccurate route choice prediction. On the other hand, the k-shortest path, link elimination, link penalty and simulation approaches generate in average a higher size of the consideration set than the historical/cohort approach. One reason is that all these approaches have explicit constrains on the maximum number of alternative paths, which is set in 20 for the deterministic methods (the k-shortest path, link elimination, link penalty approaches) and 50 for the stochastic method (simulation). Please note that only the link penalty approach always generates 20 path alternatives por each OD pair. We use the Path Size Term in Equation (7) to evaluate the capacity of each approach to generate diverse paths. PS term equal to zero means that the generated alternatives paths in the consideration set are not overlapped, then the alternatives are totally different. While the PS is smaller means that the generated alternatives paths in the consideration set are more overlapped, then the alternatives are more similar. The Simulation and Link penalty approaches generate the most heavily overlapped alternative paths, while the Historical/Cohort and Labeling approaches generate the most diverse alternative paths.

Analyzing the number of transfers per each approach, it is possible to observe that the Historical/Cohort approach generates the smallest value in average followed by the K-shortest path approach. This means that the K-shortest path approach generates, in average, fewer irrelevant path alternatives than the other approaches. A similar situation can be observed with the waiting time, in-vehicle travel time and walking transfer time, where the results obtained with the K-shortest path approach are quite similar to those of the Historical/Cohort approach. The Labeling approach generates the higher value, in average, for waiting time and the walking transfer time of the paths. This is expected as the Labeling approach applies some labels that do not take into consideration the waiting and the walking transfer times.

Table 2: Choice set generation results for each consideration set approach

Consideration set approach	Size	Path Size	N° transfers	Waiting time	In-vehicle travel time	Walking time in transfer
Historical	4.03	-1.28	0.55	7.62	26.05	1.13
K-shortest paths	7.32	-1.80	1.05	7.96	24.19	0.84
Labeling	3.86	-1.53	2.28	30.39	26.67	3.31
Link elimination	8.49	-1.60	1.39	12.89	28.80	1.39
Link penalty	20	-2.24	1.65	12.64	31.20	1.76
Simulation	9.75	-2.27	2.06	12.98	25.53	1.82

4.2 Route choice models

MNL models and PS logit models were estimated with a sample of 15,600 observations, which corresponds to trips made during 15 business days. Both types of logit models were estimated in six cases, using the Historical/Cohort, the K-shortest paths, the Labeling, the Link elimination, the Link penalty, and the Simulation approaches.

The specification of the deterministic utility function considers in-vehicle travel time, waiting time at the beginning of the trip and during transfers, walking time during transfers, and the transfer penalty, which considers bus to bus, bus to Metro, and Metro to bus transfers. Metro to Metro transfers cannot be incorporated into the model because the observed route that the passenger used inside the Metro network is not available.

Table 4 shows the estimated parameters for the MNL logit models using each consideration set generation approach. The model that uses the Historical/Cohort approach reported that all parameters are statistically significant and have the expected sign. These results are quite similar to those of models that use the Labeling, the Link elimination, the Link penalty, and the Simulation approaches, where the parameters are statistically significant and with the expected sign, except for the transfer walking time which had a positive sign in all of these heuristic approaches. In the model that uses the K-shortest path approach all the parameters where statistically significant, however, in contrast to the other heuristic approaches, it obtained the expected sign only for the transfer waiting time and the transfer parameters.

Table 5 provides the estimated parameters for the PS logit models using each consideration set generation approach. In all models the PSC term is statistically significant. Given that PSC belongs to the interval $(-\infty, 0]$ and implies a reduction in the systematic utility of correlated routes, the positive sign in the coefficient obtained for all models is expected. The models that use the Historical/Cohort, the Labeling, The Link penalty, and the Link elimination approaches maintain the results, shown in **Table 4**, in terms of statistical significance and the sign of the parameters. Different to the case of the model that uses the K-shortest paths, where the inclusion of the PSC term generated a change in the sign and significance of the parameters for in-vehicle travel time and initial waiting time, which become statistically significant and with the expected sign. In this line, the model that uses the Simulation approach also improved their results, as the parameter of transfer walking time become statistically significant and with the expected sign.

Comparing the results of the MNL models and the PS logit models, it can be seen that PS logit models has better model fit in all models. Consequently, the inclusion of the PSC term in the specification of the models allows a higher explanatory power compared to the MNL models. In summary, these results show that, if the correlation due to the overlapping between alternatives is considered in the modelling, all consideration set approaches allow representing the perception of passengers with respect to the alternative paths attributes, except for the transfer walking time attribute, which parameter obtained an unexpected sign in the Labeling, the Link penalty, the Link elimination, and the K-shortest paths approaches.

5. DISCUSSION

This section is devoted to the analysis of PS logit models parameters with the purpose of understanding the differences between each evaluated consideration set approach.

The negative coefficients of travel time, walking time, and waiting time variables show a disutility of travel time for passengers. In other words, all models indicate that alternative routes with lower in-vehicle travel time and waiting times are preferred. In the models we estimated sensitivity to waiting time at the beginning of the trip as well as at any transfer stages. The results of the rates of substitution of the initial waiting time with respect to the in-vehicle vary between 1 and 1.6. Focusing on the transfer waiting time, the rates of substitution obtained with most of the consideration set approaches, is around 2, which is a value in in line with the route choice literature on public transport (Nassir et al., 2018; S Raveau & Muñoz, 2014; Tan et al., 2015). However, the Link elimination and K-shortest paths approaches generate a rate of substitution of transfer waiting time value higher than 2, especially the model that uses the K-shortest path approach that generates a value close to 7.

The perception of passengers about the walking transfer time attribute can be explained by the models that use the Historical/Cohort and the Simulation approaches, which generate a rate of substitution around 2 for the Historical approach and 1 for the Simulation approach. The other approaches do not allow to evaluate this attribute since they generate a positive parameter for this attribute. The trade-offs between in-vehicle travel time and the walking time in transfer obtained with the Historical/Cohort and Simulation approaches are in line with the studies that have reported a value between 1 and 2, such as Jánošíkova et al. (2014) (1.7) and Nassir et al. (2018) (1.25). Raveau & Muñoz (2014) reported a value around 3 minutes for this trade-off, consequently the model that uses the Historical/Cohort approach is the one closer to this value. While Tan's work reported a value 2 times higher than the value obtained by the model that uses the Historical/Cohort approach, and 4 times higher than the value obtained by the model that uses the Simulation approach.

On the other hand, the negative coefficient of the transfer variable shows a disutility of the number of transfers for passengers. In other words, all models indicate that alternative routes with lower number of transfers are preferred. The results of the rates of substitution of the number of transfers with respect to the in-vehicle vary between 15 and 55 minutes. The smallest value is reported by the model that uses the Historical/Cohort approach and the highest values is reported by the model that uses the K-shortest path approach. Previous studies have shown that one transfer is perceived by a typical passenger as equivalent to a number that varies between 3.6 min and 16 min of in-vehicle time (Z. Guo & Wilson, 2011; Nassir et al., 2018; S Raveau & Muñoz, 2014; Tan et al., 2015). Therefore, the model that use the Historical/Cohort approach is the best representing the perception of passengers about a transfer.

For a prediction analysis, the trip dataset is split into two parts. The first part corresponds to 15,600 observations and they are used for estimation. The parameters obtained from the estimation are used to predict the path choice for the second part of the dataset, which correspond to 4,685 observations. We use the First Preference Recovery (FPR) and the Average Loglikelihood (AL) to evaluate the performance of each consideration set approach. **Table 3** shows the FPR and AL values for the models constructed with each consideration set approach, and the Historical/Cohort approach is the one with the best prediction performance.

Table 3: prediction results

Parameters	Consideration set approach					
	Historical/ Cohort	Simulation	Labeling	Link penalty	Link elimination	K-shortest paths

First preference recovery	0.5159	0.4702	0.4924	0.4608	0.4730	0.4608
Loglikelihood of validation sample	0.4362	0.2322	0.3410	0.3427	0.3444	0.2945

In summary, the results of this study suggest that the Historical/Cohort approach obtained the highest level of prediction accuracy, and it is the only one that allows to obtain, for all attributes evaluated in this study, similar rates of substitution reported in the public transport route choice modelling literature.

6. CONCLUSIONS

Modeling route choice behavior requires the identification of non-chosen paths that are considered as attractive by travelers to reach a destination. This set of alternatives is call the consideration set and it is usually unknown to the researcher in reveled preference data. The majority of the public transport route choice studies that work with this type of data have identified the consideration set through ad-hoc heuristics, such as the shortest path algorithms (Tan et al., 2015) or using historical data (Jánošíkova et al., 2014; Kim et al., 2020). We call the last approach as Historical/Cohort, which impute the observed choices in past instances of the traveler or other users in the same cohort in cross section data as the consideration set. In this study we first show the conditions under which the Historical/Cohort approach would allow recovering the population parameters. In addition, we use a case study to assess the performance of different consideration set generation approaches, which are commonly used in the public transport route choice literature, in terms of estimation and prediction. The results show that the Historical/Cohort approach surpasses the other methods in all the statistics considered.

The demonstration that proves the Historical/Cohort approach is based on an adaptation of the theorem of sampling of alternatives (McFadden, 1978), in which each historical/cohort choice is understood as a draw from the true consideration set and the sampling correction is canceled out when there are many observations. In this context, we hypothesize that the route choice models that use the Historical/Cohort approach for identifying the consideration set obtain the same or better results than any other consideration set generation approach. To evaluate this hypothesis, we use data from the public transport system in Santiago, Chile to estimate route choice models with different consideration set generation approaches: the Historical/Cohort approach, and five approaches based on shortest paths searching: the Labeling, the Link elimination, the Link penalty, the K-shortest paths, and the Simulation approaches.

We split the database in two parts, the first one corresponds to business days of three weeks and the second one corresponds to business days of one week. With the three weeks, we estimated two RUM models using each consideration set generation approach: a MNL model, which is the basic and most used model for representing route choice in public transport systems, and a PS logit model, which captures the correlation due to overlapping between alternative paths. The results show that all PS logit models obtained better fit than the MNL models, and in some cases, more robust parameter estimates. Focusing on PS logit models, the estimated parameters suggest that all consideration set generation approaches can represent the perception of passengers well for all attributes, except for the walking time in transfer, which is well represented only by the Historical/Cohort and Simulation approaches. Regarding the rates of substitution with respect to in-vehicle travel time, the Historical/Cohort approach is the only one that allows obtaining values previously reported in the public transport literature for all attributes. It is important to note that the only attribute that obtained a rate of substitution reported in previous studies with all consideration set generation approaches is the initial waiting time, for the other attributes one or more heuristic approaches reported non-expected values. The evidence from this analysis supports the idea that Historical/Cohort approach for identifying the consideration set allows estimating the population parameters. Additional to this, the comparison of prediction accuracy across different consideration set generation approaches suggest that the Cohort/Historical approach allows estimating models with better prediction abilities with respect to the choices of passengers in the prediction sample.

REFERENCES

Anderson, M. K., Nielsen, O. A., & Prato, C. G. (2017). Multimodal route choice models of public transport passengers in the Greater Copenhagen Area. *EURO Journal on Transportation and Logistics*, *6*(3), 221–245. https://doi.org/10.1007/s13676-014-0063-3

Ben-Akiva, M., Bergman, M. J., Daly, A. J., & Ramaswamy, R. (1984). Modelling inter urban route choice behaviour. Retrieved June 1, 2021, from https://trid.trb.org/view/210836

Ben-Akiva, M, & Lerman, S. (1985). *Discrete choice analysis: theory and application to travel demand*. (MIT press, Ed.) (Vol. 9). Retrieved from https://books.google.es/books?hl=es&lr=&id=7L34DwAAQBAJ&oi=fnd&pg=PR11&dq=Ben-Akiva,+M.,+y+Lerman,+S.+(1985).+Discrete+choice+analysis:+theory+and+application+to+travel+demand+(Vol.+9).+MIT+press.&ots=4U2dZN8a1R&sig=sXTEeE49N-WIUtzNucVF4wcpCF4

Ben-Akiva, Moshe. (1989). Lecture Notes of Large Set of Alternatives, with a Correction of result in BAL Chapter 9. *Unpublished Manuscript, Massachusetts Institute of Technology*.

Ben-Akiva, Moshe, & Boccara, B. (1995). Discrete choice models with latent choice sets. *International Journal of Research in Marketing*, *12*(1), 9–24. https://doi.org/10.1016/0167-8116(95)00002-J

Bliemer, M. C. J., & Bovy, P. H. L. (2008). Impact of Route Choice Set on Route Choice Probabilities. *Transportation Research Record: Journal of the Transportation Research Board*, *2076*(1), 10–19. https://doi.org/10.3141/2076-02

Bovy, P. H. L., Bekhor, S., & Prato, C. G. (2008). The Factor of Revisited Path Size. *Transportation Research Record: Journal of the Transportation Research Board*, *2076*(1), 132–140. https://doi.org/10.3141/2076-15

Bovy, P. H. L., & Hoogendoorn-Lanser, S. (2005). Modelling route choice behaviour in multi-modal transport networks. *Transportation*, *32*(4), 341–368. https://doi.org/10.1007/s11116-004-7963-2

Brown, J. J., & Wildt, A. R. (1992). Consideration Set Measurement. *Journal of the Academy of Marketing Science*, *20*(3), 235–243.

de Grange, L., Raveau, S., & González, F. (2012). A Fixed Point Route Choice Model for Transit Networks that Addresses Route Correlation. *Procedia - Social and Behavioral Sciences*, *54*, 1197–1204. https://doi.org/10.1016/j.sbspro.2012.09.834

Eluru, N., Chakour, V., & El-Geneidy, A. M. (2012). Travel mode choice and transit route choice behavior in Montreal: Insights from McGill University members commute patterns. *Public Transport*, *4*(2), 129–149. https://doi.org/10.1007/s12469-012-0056-2

Fiorenzo-Catalano, S., Van Nes, R., & Bovy, P. H. L. (2004). *Choice Set Generation for Multi-modal Travel Analysis*. *journals.open.tudelft.nl*. Retrieved from https://journals.open.tudelft.nl/ejtir/article/view/4262

Grison, E., Burkhardt, J.-M., & Gyselinck, V. (2017). How do users choose their routes in public transport? The effect of individual profile and contextual factors. *Transportation Research Part F: Traffic Psychology and Behaviour*, *51*, 24–37. https://doi.org/10.1016/J.TRF.2017.08.011

Gschwender, A., Munizaga, M., & Simonetti, C. (2016). Using smart card and GPS data for policy and planning: The case of Transantiago. *Research in Transportation Economics*, *59*, 242–249. https://doi.org/10.1016/J.RETREC.2016.05.004

Guevara, C. A., & Ben-Akiva, M. E. (2013a). Sampling of alternatives in Logit Mixture models. *Transportation Research Part B: Methodological*, *58*, 185–198. https://doi.org/10.1016/j.trb.2013.08.011

Guevara, C. A., & Ben-Akiva, M. E. (2013b). Sampling of alternatives in Multivariate Extreme Value (MEV) models. *Transportation Research Part B: Methodological*, *48*, 31–52. https://doi.org/10.1016/j.trb.2012.11.001

Guevara, C. A., Chorus, C. G., & Ben-Akiva, M. E. (2016). Sampling of alternatives in random regret minimization models. *Transportation Science*, *50*(1), 306–321. https://doi.org/10.1287/trsc.2014.0573

Guo, S., Yu, L., Chen, X., & Zhang, Y. (2011). Modelling waiting time for passengers transferring from rail to buses. *Transportation Planning and Technology*, *34*(8), 795–809. https://doi.org/10.1080/03081060.2011.613589

Guo, Z. (2011). Mind the map! The impact of transit maps on path choice in public transit. *Transportation Research Part A: Policy and Practice*, *45*(7), 625–639. https://doi.org/10.1016/j.tra.2011.04.001

Guo, Z., & Wilson, N. H. M. (2011). Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground. *Transportation Research Part A: Policy and Practice*, *45*(2), 91–104. https://doi.org/10.1016/j.tra.2010.11.002

Hauser, J. R. (2014). Consideration-set heuristics. *Journal of Business Research*, *67*(8), 1688–1699. https://doi.org/10.1016/j.jbusres.2014.02.015

Hoogendoorn-Lanser, S., & Van Nes, R. (2004). Multimodal Choice Set Composition:

Analysis of Reported and Generated Choice Sets. *Transportation Research Record: Journal of the Transportation Research Board*, *1898*(1), 79–86. https://doi.org/10.3141/1898-10

Hoogendoorn-Lanser, S., van Nes, R., & Bovy, P. (2005). Path Size Modeling in Multimodal Route Choice Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, *1921*(1), 27–34. https://doi.org/10.1177/0361198105192100104

Hoogendoorn-Lanser, S., Van Nes, R., & Hoogendoorn, S. P. (2006). Modeling Transfers in Multimodal Trips. *Transportation Research Record: Journal of the Transportation Research Board*, *1985*(1), 144–153. https://doi.org/10.1177/0361198106198500116

Jánošíkova, L., Slavík, J., & Koháni, M. (2014). Estimation of a route choice model for urban public transport using smart card data. *Transportation Planning and Technology*, *37*(7), 638–648. https://doi.org/10.1080/03081060.2014.935570

Kim, I., Kim, H.-C., Seo, D.-J., Jung, ·, & Kim, I. (2020). Calibration of a transit route choice model using revealed population data of smartcard in a multimodal transit network. *Transportation*, *47*, 2179–2202. https://doi.org/10.1007/s11116-019-10008-8

Manski C. F. (1977). The structure of random utility models. *Theory and Decision*, *8*(3), 229–254. Retrieved from https://search.proquest.com/openview/7acf07ef00e4d7b837b4de87994aed40/1?pq-origsite=gscholar&cbl=1818302

McFadden, D. (1978). "Modeling the Choice of Residential Location," in Karlquist, Lundqvist, Snickers and Weibull eds. *Spatial Interaction Theory and Residential Location, North Holland, Amsterdam*, 75–96.

Mora-Garcia, R. T., Marti-Ciriquian, P., Perez-Sanchez, R., & Cespedes-Lopez, M. F. (2018). A comparative analysis of manhattan, euclidean and network distances. Why are network distances more useful to urban professionals? In *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management, SGEM* (Vol. 18, pp. 3–10). https://doi.org/10.5593/sgem2018/2.2/S08.001

Munizaga, M. A., & Palma, C. (2012). Estimation of a disaggregate multimodal public transport Origin–Destination matrix from passive smartcard data from Santiago, Chile. *Transportation Research Part C: Emerging Technologies*, *24*, 9–18. https://doi.org/10.1016/J.TRC.2012.01.007

Nassir, N., Hickman, M., & Ma, Z.-L. (2018). A strategy-based recursive path choice model for public transit smart card data. *Transportation Research Part B: Methodological*, 1–21. https://doi.org/10.1016/j.trb.2018.01.002

Pelletier, M., Trépanier, M., C, C. M.-T. R. P., & 2011, undefined. (n.d.). Smart card data use in public transit: A literature review. *Elsevier*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0968090X1000166X?casa_token=RQONoNPqvRgAAAAA:2T5dzQ1S4jQyHk0PxBTr16QDe1Kr7h7mqSET-l0U-wYPTwYnK08tZ_GC1EUmyQMzxlwvyw556f5U

Prato, C. G. (2009). Route choice modeling: Past, present and future research directions. *Journal of Choice Modelling*, *2*(1), 65–100. https://doi.org/10.1016/S1755-5345(13)70005-8

Prato, C. G., & Bekhor, S. (2007). Modeling Route Choice Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, *2003*(1), 64–73. https://doi.org/10.3141/2003-09

Raveau, S, & Muñoz, J. C. (2014). Analyzing route choice strategies on transit networks. Retrieved from https://trid.trb.org/view/1289224

Raveau, Sebastián, Guo, Z., Muñoz, J. C., & Wilson, N. H. M. (2014). A behavioural comparison of route choice on metro networks: Time, transfers, crowding, topology and socio-demographics. *Transportation Research Part A: Policy and Practice*, *66*(1), 185–195. https://doi.org/10.1016/j.tra.2014.05.010

Raveau, Sebastián, Muñoz, J. C., & de Grange, L. (2011). A topological route choice model for metro. *Transportation Research Part A: Policy and Practice*, *45*(2), 138–147. https://doi.org/10.1016/J.TRA.2010.12.004

Rui, T. (2016). *Modeling Route Choice Behaviour in Public Transport Network*. *National Universty of Singapore*. Retrieved from https://core.ac.uk/download/pdf/83107872.pdf

Swait, J., & Ben-Akiva, M. (1987). Incorporating random constraints in discrete models of choice set generation. *Transportation Research Part B: Methodological*, *21*(2), 91–102. Retrieved from https://www.sciencedirect.com/science/article/pii/0191261587900099

Tan, R., Adnan, M., Lee, D.-H., & Ben-Akiva, M. E. (2015). New Path Size Formulation in Path Size Logit for Route Choice Modeling in Public Transport Networks. *Transportation Research Record: Journal of the Transportation Research Board*, *2538*(1), 11–18. https://doi.org/10.3141/2538-02

Tien, D. N., MacDonald, T., & Xu, Z. (2011). TDplanner: Public transport planning system with real-time route updates based on service delays and location tracking. In *IEEE Vehicular Technology Conference*. https://doi.org/10.1109/VETECS.2011.5956479

Ton, D., Shelat, S., Nijënstein, S., Rijsman, L., van Oort, N., & Hoogendoorn, S. (2020). Understanding the Role of Cycling to Urban Transit Stations through a Simultaneous Access Mode and Station Choice Model. *Transportation Research Record*, *2674*(8), 823–835. https://doi.org/10.1177/0361198120925076

Villalobos, N., & Guevara, C. A. (2021). Caracterización del conjunto de consideración en elección de ruta. *Estudios de Transporte*, *22*(1), 1–26. Retrieved from https://estudiosdetransporte.org/sochitran/article/view/247

Vrtic, M. ;, & Axhausen, K. W. (2002). The impact of tilting trains in Switzerland: A route choice model of regional-and long distance public transport trips paper submitted to the 82nd Annual Meeting of the Transportation Research Board. https://doi.org/10.3929/ethz-a-004403563

Table 4: MNL model estimates (t tests) using each generation consideration set approach

Parameters	Consideration set approach					
	Historical/Cohort	**Simulation**	**Labeling**	**Link penalty**	**Link elimination**	**K-shortest paths**
Travel time in vehicle	-0.131* (-44.1)	-0.073* (-24.1)	-0.167* (-58.6)	-0.145* (-80.7)	-0.115* (-56.4)	0.007 (1.7)
Initial waiting time	-0.173* (-39.6)	-0.079* (-17.1)	-0.106* (-24.2)	-0.205* (-49.4)	-0.164* (-38.5)	0.036* (6.9)
Transfer waiting time	-0.266* (-20.5)	-0.340* (-34.0)	-0.248* (-20.9)	-0.350* (-38.6)	-0.342* (-29.1)	-0.436* (-25.2)
Transfer walking time	-0.217* (-19.0)	0.110* (8.6)	0.429* (23.0)	0.173* (14.8)	0.428* (31.6)	0.493* (31.1)
Transfer	-1.448* (-17.3)	-2.267* (-33.5)	-2.12* (-26.8)	-3.535* (-58.5)	-2.75* (-39.3)	-2.829* (-35.8)
N° observations	15,600	15,600	15,600	15,600	15,600	15,600
Log-likelihood	-17,169.6	-14,597.1	-14,866.5	-21,207.4	-17,445.8	-14503.6
Adjusted rho-square	0.129	0.522	0.287	0.5467	0.4443	0.3814
AIC	34,349.2	29,204.1	29,742.9	42,424.8	34,901.5	29,017.2

All columns show t-values between parentheses. AIC = Akaike information criterion. *significant on the 5% level.

Table 5: PS Logit model estimates (t tests) using each generation consideration set approach

Parameters	Consideration set approach					
	Historical/Cohort	**Simulation**	**Labeling**	**Link penalty**	**Link elimination**	**K-shortest paths**
Travel time in vehicle	-0.103* (-33.0)	-0.157* (-34.9)	-0.132* (-39.7)	-0.136* (-75.3)	-0.104* (-46.6)	-0.059* (-12.1)
Initial waiting time	-0.142* (-32.0)	-0.199* (-32.8)	-0.134* (-25.4)	-0.218* (-51.3)	-0.150* (-33.3)	-0.063* (-11.1)
Transfer waiting time	-0.217* (-15.8)	-0.286* (-24.0)	-0.247* (-17.4)	-0.303* (-32.5)	-0.320* (-24.6)	-0.397* (-19.2)
Transfer walking time	-0.198* (-17.4)	-0.162* (-6.8)	0.182* (6.1)	0.046* (3.1)	0.267* (16.6)	0.321* (17.2)
Transfer	-1.583* (-18.7)	-3.442* (-30.8)	-2.638* (-26.5)	-3.877* (-60.2)	-2.667* (-35.1)	-3.182* (-34.5)
PSC	1.009* (27.5)	1.851* (59.6)	1.813* (67.3)	1.064* (61.4)	0.983* (57.1)	1.270* (52.5)
N° observations	15,600	15,600	15,600	15,600	15,600	15,600
Log-likelihood	-16,699.2	-9,201.4	-10,988.8	-19,756.0	-16,002.8	-13,334.4
Adjusted rho-square	0.152	0.699	0.473	0.5777	0.4903	0.4312
AIC	33,410.3	18,414.9	21,989.7	39,524.0	32,017.6	26680.8

All columns show t-values between parentheses. AIC = Akaike information criterion. *significant on the 5% level